Investigation of False Learning in Adaptive Monte Carlo Transport

Henry Lichtenstein

August 15, 1996

PRESENTATION OUTLINE

- **Introduction**
- **Preliminaries**
- **Demonstration of False Learning**
- **An Avoidance Strategy**
- **Diagnostics**
- **Summary**

INTRODUCTION

False Learning (FL) in Adaptive Monte Carlo:

Learning falsely **that a domain of phase space is a relatively unimportant contributor to the result; due to adaptively inadequate sampling. FL can lead to** ***false convergence*** **(i.e.,** ***convergence*** **to a** ***false*** **result).**

Demonstration of FL:

Contrive a transport problem that poses a potential FL situation, which requires ***heroic a priori measures*** **(i.e., ultra conservative initial sampling) to avoid false convergence.**

An Avoidance Strategy:

Strive **to insure that no sequence of transitions, which could produce significant score, is ignored.**

FL Diagnostics:

Computed feedback that, based on a comparison of known and estimated information, will suggest the presence of FL in the course of the calculation.

PRESENTATION OUTLINE

- **Introduction**

  Preliminaries

- **Demonstration of False Learning**
- **An Avoidance Strategy**
- **Diagnostics**
- **Summary**

PRELIMINARIES

Testbed for Development of Algorithms:

A multi-state discrete Monte Carlo transport code, written previously by Tom Booth. Code adaptively iterates to zero-variance biasing:

$q_{ij} = C_i * p_{ij} * (s_{ij} + m_j)$; *transition* from state_i to state_j;

(for *termination* state_0, $m_0 = 0$);

where

q = zero-variance biased transition (termination) probability;

p = unbiased transition (termination) probability;

s = corresponding transition (termination) score;

m = estimated mean;

C = normalization.

PRELIMINARIES (continued)

Reference (Benchmark) for Development of Algorithms:

Wrote code that iteratively computes analytic values of means for corresponding multi-state discrete Monte Carlo transport scenario, based on:

mean_*i* = SUMMATION_*(j=0,n)* {p_*ij{s_*ij* + mean_*j*}}**

where *n* is number of states and state_*0* is termination (mean_*0* is 0).

Code computes analytic means within machine *double* precision, using internally computed criterion, *viz.*

```
      subroutine epsil
      implicit double precision (a-h,o-z)
      common/epsilon/eps
c------ compute machine epsilon;
      eps=1.
      do 10 i=1,1000
      if(1.0+eps.eq.1.0)go to 20
10    eps=eps/2.0
20    eps=2.0*eps
      write(*,*)'machine epsilon=',eps
      return
      end
```

PRESENTATION OUTLINE

- **Introduction**
- **Preliminaries**
- ➡ **Demonstration of False Learning**
- ○ **An Avoidance Strategy**
- ○ **Diagnostics**
- ○ **Summary**

FALSE LEARNING DEMONSTRATION

Zero-variance biased transition probabilities (beyond initial iteration):

$$q_{ij} = C_i * p_{ij} * (s_{ij} + m_j); \quad (m_0 = 0)$$

Contrive 2-state problem (source in state_*1*), with following (transition probabilities; associated scores):

	j=0	j=1	j=2
i=1	(.990; 1.0)	(.009; 1.0)	(.001; 0.0)
i=2	(.990; 0.0)	(.001; 1e6)	(.009; 0.0)

With un-biased first iteration, compute:

Initial histories	Final m_1	Final m_2	False learning	Convergence
1,000,000 (*insufficient!*)	1.0081 (.0001)	(*no estimate!*)	initial_m_2=0 --> q_{12}=0	FALSE
10,000,000 (*heroic?*)	2.0263 (.0003)	1009.1 (.1)	AVOIDED	PROPER
analytic values	2.02632066 ...	1009.08378 ...	N/A	N/A

PRESENTATION OUTLINE

- **Introduction**
- **Preliminaries**
- **Demonstration of False Learning**
- **An Avoidance Strategy**
- **Diagnostics**
- **Summary**

An AVOIDANCE STRATEGY

Rationale:

- **Need to inspect *adequately* all states that contribute significant score.**
- **In general, adequate inspection of all states is impractical.**
- **A strategy that *tends* to inspect all states may be a useful first step.**
- **Given a finite first iteration, how can unbiased probabilities be modified to have such tendency?**

Test Strategy:

- **Modify Booth's multi-state code to allow biasing for first iteration.**
- **Investigate *uniform*-biased *a priori* transition probabilities.**

An AVOIDANCE STRATEGY (continued)

Test Scenario:

- **Restrict transitions to a contiguous state or termination only, with source in state_*1*.**
- **Disallowing other (including in-state) transitions merely accelerates testing.**
- **Assign 0.1 probability for all state-to-state transitions.**
- **Assign remaining probability to termination (0.9 for "end" state or 0.8 for "interior" state).**
- **For source state_*1*, assign score of 1.0 for termination; 0.0 score for transition to state_*2*.**
- **For all states_*j*, *j* > *1*, assign score of 100.0 for "downscattering" to state_*j-1*; 0.0 score otherwise.**

Medium is relatively opaque to arrival at higher states.
Higher states contribute score, hence difficult FL avoidance.

An AVOIDANCE STRATEGY (continued)

An Example Scenario:

For a 6-state scenario having prescribed (transition; scoring) properties, state-to-state matrix is:

	j=0	j=1	j=2	j=3	j=4	j=5	j=6
i=1	(.9; 1.)	(.0; 0.)	(.1; 0.)	(.0; 0.)	(.0; 0.)	(.0; 0.)	(.0; 0.)
i=2	(.8; 0.)	(.1; 100.)	(.0; 0.)	(.1; 0.)	(.0; 0.)	(.0; 0.)	(.0; 0.)
i=3	(.8; 0.)	(.0; 0.)	(.1; 100.)	(.0; 0.)	(.1; 0.)	(.0; 0.)	(.0; 0.)
i=4	(.8; 0.)	(.0; 0.)	(.0; 0.)	(.1; 100.)	(.0; 0.)	(.1; 0.)	(.0; 0.)
i=5	(.8; 0.)	(.0; 0.)	(.0; 0.)	(.0; 0.)	(.1; 100.)	(.0; 0.)	(.1; 0.)
i=6	(.9; 0.)	(.0; 0.)	(.0; 0.)	(.0; 0.)	(.0; 0.)	(.1; 100.)	(.0; 0.)

source in state_*1*; termination is state_*0*

''Conditions of contest'' for investigation:

1. **Seek threshold number of histories required in first (*i.e., a priori* learning) iteration to compute *non-zero* mean for every state.**
2. **Incrementing algorithm: try 10x histories every time *known* non-zero mean is computed as zero.**

An AVOIDANCE STRATEGY (continued)

Comparison of threshold number of histories in initial iteration (*i.e., a priori* learning stage) required to *avoid learning falsely* that a known *non*-zero mean is zero:

# of States	Threshold histories for *a priori* un-biased	Threshold histories for *a priori* uniform-biased
2	100	10
3	1,000	10
4	100,000	100
5	100,000	100
6	10,000,000	1,000

Conjecture:
Tendency to inspect all states has *tendency* to avoid FL.

PRESENTATION OUTLINE

- **Introduction**
- **Preliminaries**
- **Demonstration of False Learning**
- **An Avoidance Strategy**
- **Diagnostics**
- **Summary**

DIAGNOSTICS

Desired Attributes for FL Diagnostic Flag:

- ***True* indicator of FL presence;**
  (tautology, but important to note);
- **Succinct (for easy comprehension); one number or small set (such as MCNP statistical tests);**
- **Computable without undue diversion of resources from main calculation;**
- ***Ideally*, would *appear* in FL presence; *vanish* in FL absence; *less-than-ideally*, diminished magnitude in FL absence;**

Recall that FL is introduced when undersample scoring-transitions; FL enhanced when undersampled-scoring relatively large. This suggests basis for FL flag:

> **Comparison between *computed* and *theoretical* values of state-to-state branching weight, perhaps *weighted* by associated scores.**

Suggested basis for FL flag recognizes that, although the global solution is unknown, local behavior is known.

DIAGNOSTICS (continued)

Initial Definition of an FL Diagnostic Flag:

Let S_c be the *estimated* quantity

$$S_c = \mathrm{SUM}_{(i=1,n)}\ \mathrm{SUM}_{(j=0,n)}\ \{(b_{ij})*(s_{ij})/(w_i)\}\ ;$$

let S_t be the *known* quantity

$$S_t = \mathrm{SUM}_{(i=1,n)}\ \mathrm{SUM}_{(j=0,n)}\ \{(p_{ij})*(s_{ij})\}\ .$$

Then define FL flag to be:

$$\mathrm{flag} = 1 - S_c/S_t$$

where n is number of discrete states ($state_0$ is termination); w_i is total weight entering $state_i$; b_{ij} is weight branching from $state_i$ to $state_j$; p_{ij} is transition probability from $state_i$ to $state_j$; s_{ij} is associated score; and *i.ne.k* for $w_k = 0$. Note flag is single computed quantity, independent of n.

Flag vanishes for $S_c = S_t$, which is unlikely in presence of FL, for small # of non-zero s_{ij}.

DIAGNOSTICS (continued)

Test of Initial Definition for an FL Flag

Test Scenario: tight coupling between contiguous states; source in state_*1*; termination prob = .4 in all states; transition (to nearest neighbor) prob = .3 (.6 for end states); score = 1 for termination in highest state (state_*n*, where *n* varies from *2* to *8* in this test). Machine epsilon, *eps* = 2.2e-16; state_*m* has max %_rel.

n_states	state_*1*	(%_rel)	state_*m*	(%_rel)	state_*n*	(%_rel)	(flag)
2	0.37500	(.005)	0.62500	(.01)	0.62500	(.01)	(+ .02)
	0.374999 ...	*eps*	0.624999 ...	*eps*	0.624999 ...	*eps*	
3	0.11250	(.01)	0.51251	(.02)	0.51251	(.02)	(- .02)
	0.112499 ...	*eps*	0.512499 ...	*eps*	0.512499 ...	*eps*	
4	0.037088	(.009)	0.16896	(.01)	0.50137	(.004)	(+ .01)
	0.0370879 ...	*eps*	0.168956 ...	*eps*	0.501373 ...	*eps*	
5	0.012347	(.01)	0.012347	(.01)	0.50015	(.002)	(- .0003)
	0.0123475 ...	*eps*	0.0123475 ...	*eps*	0.500152 ...	*eps*	
6	0.0041153	(.008)	0.55633	(.01)	0.50002	(.01)	(- .02)
	0.00411529 ...	*eps*	0.556327 ...	*eps*	0.500016 ...	*eps*	
7	0.0013717	(.009)	0.62490	(.02)	0.50000	(.003)	(+ .01)
	0.00137174 ...	*eps*	0.624905 ...	*eps*	0.500000 ...	*eps*	
8	0.00045725	(.008)	0.61813	(.01)	0.500000	(.002)	(- .02)
	0.000457250 ...	*eps*	0.618130 ...	*eps*	0.5000000 ...	*eps*	

DIAGNOSTICS (continued)

Test of Initial Definition for an FL Flag (continued)

Modified Test Scenario: same as initial, except for state_*n*, which has termination prob = .999 and transition (to lower neighbor) prob = .001; score for transition = 400. Encourages FL.

n_states	state_*1*	(%_rel)	state_*m*	(%_rel)	state_*n*	(%_rel)	(flag)
2	0.40000	(.02)	0.40000	(.02)	*no estimate*	*N/A*	(+ .5)
	0.640384 ...	*eps*	0.640384 ...	*eps*	0.400640 ...	*eps*	
3	0.48780	(.02)	0.14634	(.01)	*no estimate*	*N/A*	(+ .5)
	0.575674 ...	*eps*	0.292790 ...	*eps*	0.400292 ...	*eps*	
4	0.49863	(.005)	0.16438	(.01)	*no estimate*	*N/A*	(+ .5)
	0.528232 ...	*eps*	0.213721 ...	*eps*	0.400184 ...	*eps*	
5	0.49984	(.003)	0.548612	(.01)	*no estimate*	*N/A*	(+ .5)
	0.509726 ...	*eps*	0.9986447 ...	*eps*	0.4001500 ...	*eps*	
6	0.500000	(.0006)	0.16663	(.01)	*no estimate*	*N/A*	(+ .5)
	0.5032763 ...	*eps*	0.172127 ...	*eps*	0.400138 ...	*eps*	
7	0.500000	(.0003)	0.6098	(.03)	*no estimate*	*N/A*	(+ .5)
	0.5010958 ...	*eps*	0.50562 ...	*eps*	0.400135 ...	*eps*	
8	0.5000002	(.00006)	0.18296	(.01)	*no estimate*	*N/A*	(+ .5)
	0.50036571 ...	*eps*	0.046492 ...	*eps*	0.400133 ...	*eps*	

DIAGNOSTICS (continued)

Test of Initial Definition for an FL Flag (continued)

Modified Test Scenario: same as preceding, except use *a priori* uniform biasing to avoid FL.

n_states	state_*1*	(%_rel)	state_*m*	(%_rel)	state_*n*	(%_rel)	(flag)
2	0.64038	(.01)	0.64038	(.01)	0.40064	(.009)	(- .002)
	0.640384 ...	*eps*	0.640384 ...	*eps*	0.400640 ...	*eps*	
3	0.57568	(.01)	0.29280	(.01)	0.40029	(.01)	(- .001)
	0.575674 ...	*eps*	0.292790 ...	*eps*	0.400292 ...	*eps*	
4	0.52823	(.01)	0.52823	(.01)	0.40018	(.01)	(+ .001)
	0.528232 ...	*eps*	0.528232 ...	*eps*	0.400184 ...	*eps*	
5	0.50973	(.006)	0.40015	(.02)	0.40015	(.02)	(- .002)
	0.509726 ...	*eps*	0.400150 ...	*eps*	0.400150 ...	*eps*	
6	0.50327	(.009)	0.40014	(.01)	0.40014	(.01)	(+ .006)
	0.503276 ...	*eps*	0.400138	*eps*	0.400138 ...	*eps*	
7	0.50105	(.01)	0.40014	(.02)	0.40014	(.02)	(- .002)
	0.501095 ...	*eps*	0.400135 ...	*eps*	0.400135 ...	*eps*	
8	0.50036	(.004)	0.020985	(.02)	0.40013	(.01)	(- .003)
	0.500365 ...	*eps*	0.0209864 ...	*eps*	0.400133 ...	*eps*	

FL DIAGNOSTICS (continued)

FOLLOW-ON WORK

FL diagnostics, based on estimates of cumulative branching fractions, will probably prove inadequate for large numbers of scoring states:

- **Search for alternative flag-basis that can function well, independent of number of states.**

- **Extend diagnostics algorithms to continuous transport scenarios.**

PRESENTATION OUTLINE

- **Introduction**
- **Preliminaries**
- **Demonstration of False Learning**
- **An Avoidance Strategy**
- **Diagnostics**
- **Summary**

SUMMARY

- **Straightforward to contrive and demonstrate FL for zero-variance biasing:**

$$q_{ij} = C_i * p_{ij} * (s_{ij} + m_j)$$

1. **Specify a significant score, *i.e.*, $s_{jk} > 0$,**
2. **for one low-probability transition from state_*j*, *i.e.*, $p_{jk} \ll 1$,**
3. **to effect initial-estimate of m_j as 0,**
4. **(from that phase-space region, *i.e.*, state_*j*),**
5. **which has no associated-score upon entry from state_*i*, *i.e.*, $s_{ij} = 0$,**
6. **whereby q_{ij} set to 0 (= FL), which leads to false convergence.**

- **Some evidence that *tendency* to inspect all states has *tendency* to avoid FL.**

- **Basis for FL diagnostics is recognition that, although *global solution* is unknown, *local behavior* is known.**